**AP Stat- Worksheet 8A Linear Regression: AIRFARES**

NEED: prgrm CORR, group AIRFARES (lists DIST and AIRF)

The lists DIST and AIRF contain the following data:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Destination** | **Distance** | **Airfare** | **Destination** | **Distance** | **Airfare** |
| Atlanta | 576 | 178 | Miami | 946 | 198 |
| Boston | 370 | 138 | New Orleans | 998 | 188 |
| Chicago | 612 | 94 | New York | 189 | 98 |
| Dallas | 1216 | 278 | Orlando | 787 | 179 |
| Detroit | 409 | 158 | Pittsburgh | 210 | 138 |
| Denver | 1502 | 258 | St. Louis | 737 | 98 |

1. Create a scatterplot of distance vs. airfare and draw it below. Describe the form, direction and strength. Do there appear to be any outliers?
2. Use the program CORR to compute the mean and standard deviation of distance and airfare, and the value of the correlation between the two. Record the values below.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Mean** | **Std. Dev.** | **Correlation** |
| **Distance** |  |  |  |
| **Airfare** |  |  |  |

1. ***Use these statistics and the AP formulas*** to calculate the least squares coefficients (write your final answer as *y=b0 + b1x*)
2. Find the equation of the LSR line using the calculator. Do all the values agree with your calculations in #3? Add the line to your scatterplot in #1. Also, find the correlation coefficient and state it below.
3. What airfare does the LSR line predict for a destination that is 370 miles away? What about for 1502 miles? Using the actual data given, calculate the residual for each of these estimates (*residual = actual value – predicted value*).
4. What airfare would the regression line predict for a flight to San Francisco, which is 2842 miles from Baltimore? Would you take this prediction as seriously as the ones above? Why or why not?
5. The actual airfare for SF at that time was $198. This illustrates the danger of ***extrapolation*** *(predicting outside our x-values range)*. Give the residual for your prediction to SF.
6. Use the equation of the regression line to predict the airfare if the distance is 900 miles. Record the prediction below, and fill in the rest of the table.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| distance | 900 | 901 | 902 | 903 | 904 | 905 |
| airfare |  |  |  |  |  |  |

1. Do you notice a pattern in these predictions? By how many dollars is each prediction higher than the preceding one? Does this number look familiar (form your earlier calculations)? Explain.
2. This shows how we can interpret the slope of the LSR line. Interpret the slope in context of the problem.
3. By how much does the regression line predict airfare to rise for each additional 100 miles that a destination is farther away?